

Tidyverse Problem Set

MA 615

Yiru Fei

10/2/2019

Problem 1

Load the gapminder data from the gapminder package.

How many continents are included in the data set?

There are 5 continents in the data set.

How many countrys are included? How many countries per continent?

There are 142 countries. Africa has 52 countries, Americas has 25 countries, Asia has 33 countries, Europe has 30 countries and Oceania has 2 countries.

Using the gapminder data, produce a report showing the continents in the dataset, total population per continent, and GDP per capita. Be sure that the table is properly labeled and suitable for inclusion in a printed report.

Table 1: Total population and total GDP for each continents

continent	population_million	GDP_million
Africa	6187.5860	1.3689029
Americas	7351.4385	2.1408331
Asia	30507.3339	3.1292516
Europe	6181.1153	5.2090112
Oceania	212.9921	0.4469186

Produce a well-labeled table that summarizes GDP per capita for the countries in each continent, contrasting the years 1952 and 2007.

Table 2: Summary GDP for each continents in 1952

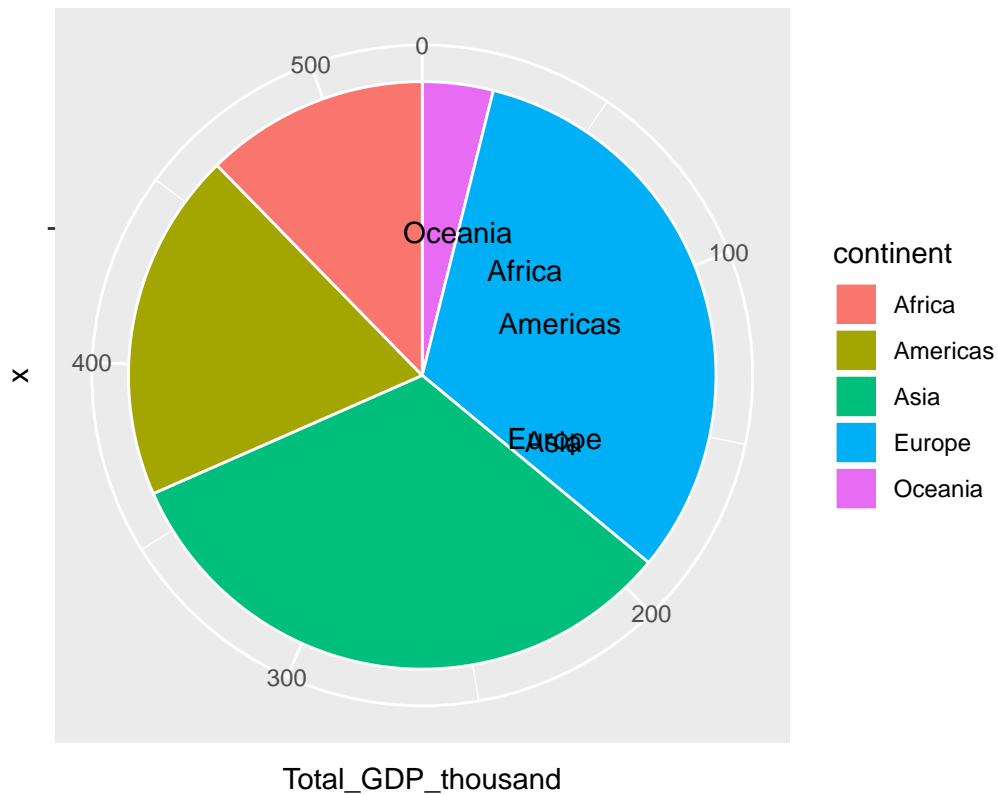
continent	Total_GDP_thousand	Ave_GDP_thousand	Max_GDP_thousand	Min_GDP_thousand
Africa	65.13377	1.252573	4.725295	0.2988462
Americas	101.97656	4.079063	13.990482	1.3977171
Asia	171.45097	5.195484	108.382353	0.3310000
Europe	169.83172	5.661057	14.734233	0.9735332
Oceania	20.59617	10.298086	10.556576	10.0395956

Table 3: Summary GDP for each continents in 2007

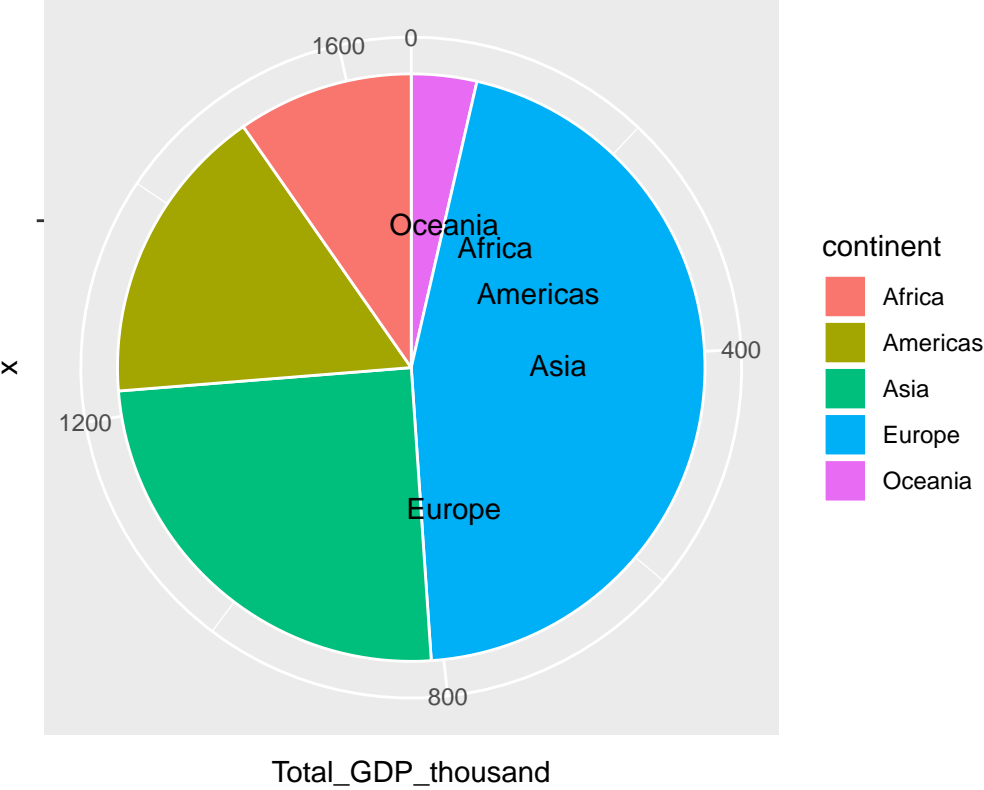
continent	Total_GDP_thousand	Ave_GDP_thousand	Max_GDP_thousand	Min_GDP_thousand
Africa	160.62970	3.089033	13.20648	0.2775519
Americas	275.07579	11.003032	42.95165	1.2016372
Asia	411.60989	12.473027	47.30699	0.9440000
Europe	751.63445	25.054482	49.35719	5.9370295
Oceania	59.62038	29.810188	34.43537	25.1850091

Product a plot that summarizes the same data as the table. There should be two plots per continent.

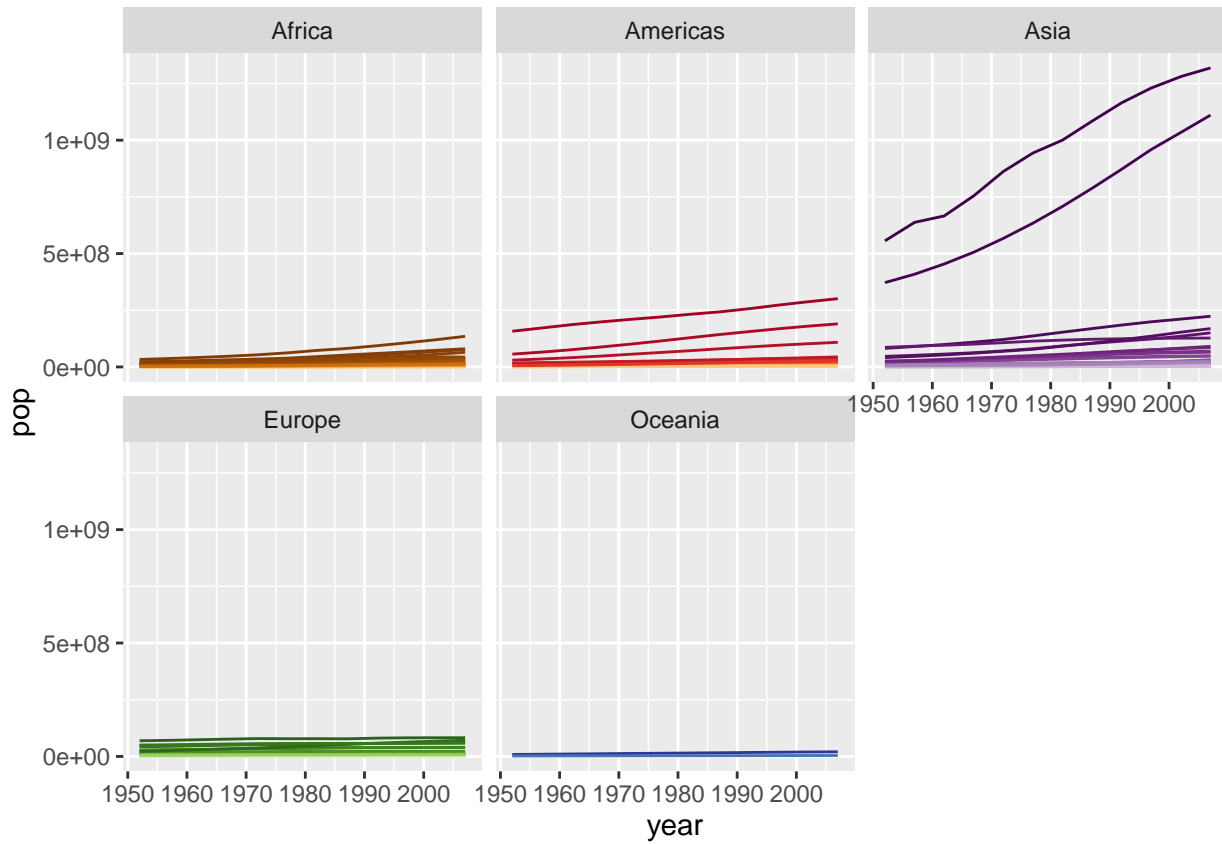
Total GDP of each continent in 1952



Total GDP of each continent in 2007



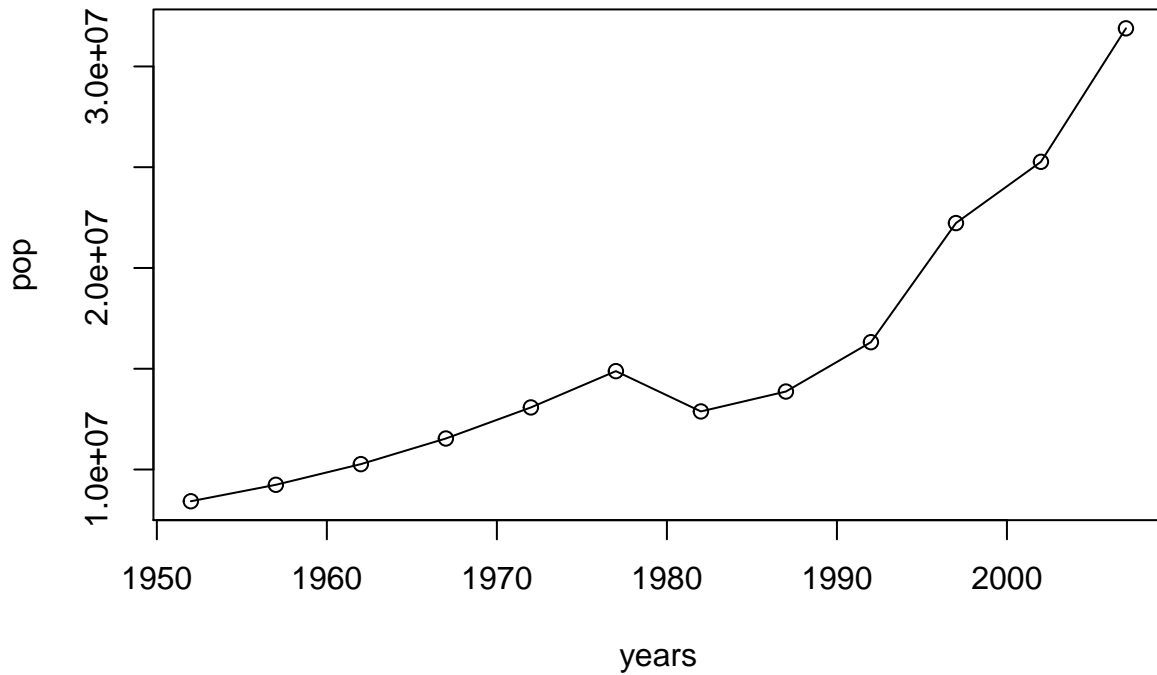
Which countries in the dataset have had periods of negative population growth?
Illustrate your answer with a table or plot.



Asia had periods of negative population growth.

For Afghanistan:

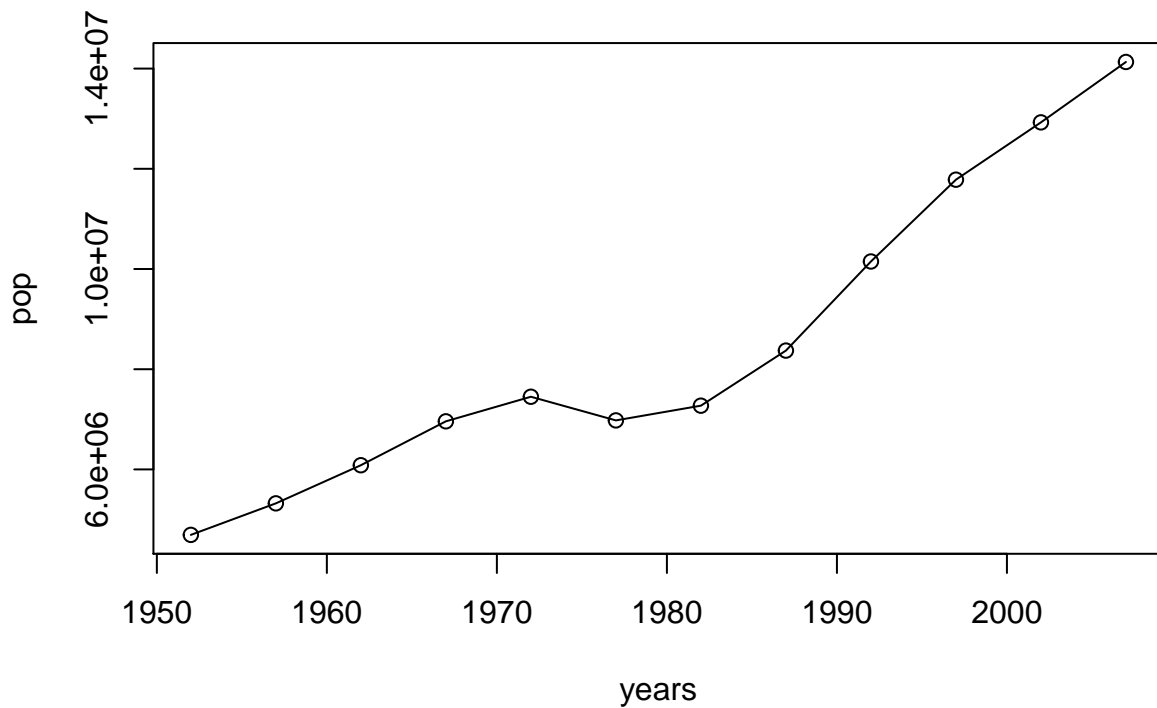
Total population in Afghanistan from 1952 to 2007



There was a decrease from 1977 to 1982 in Afghanistan.

For Cambodia:

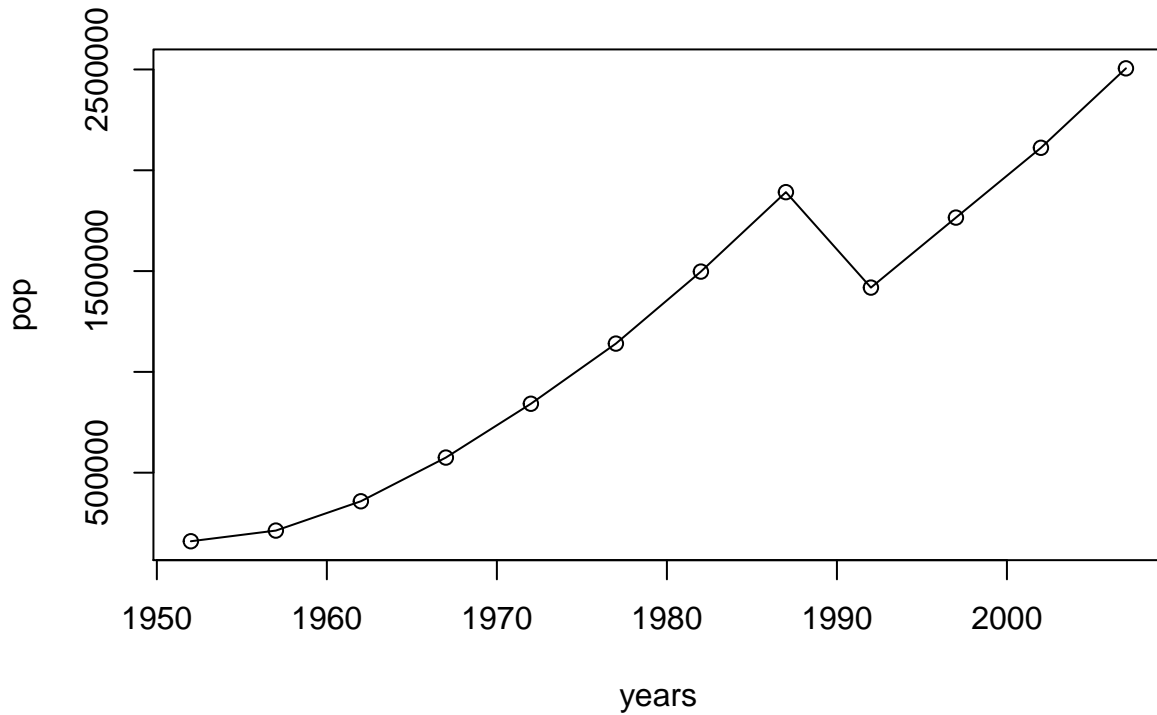
Total population in Afghanistan from 1952 to 2007



There was a decrease from 1972 to 1977 in Cambodia.

For Kuwait:

Total population in Afghanistan from 1952 to 2007

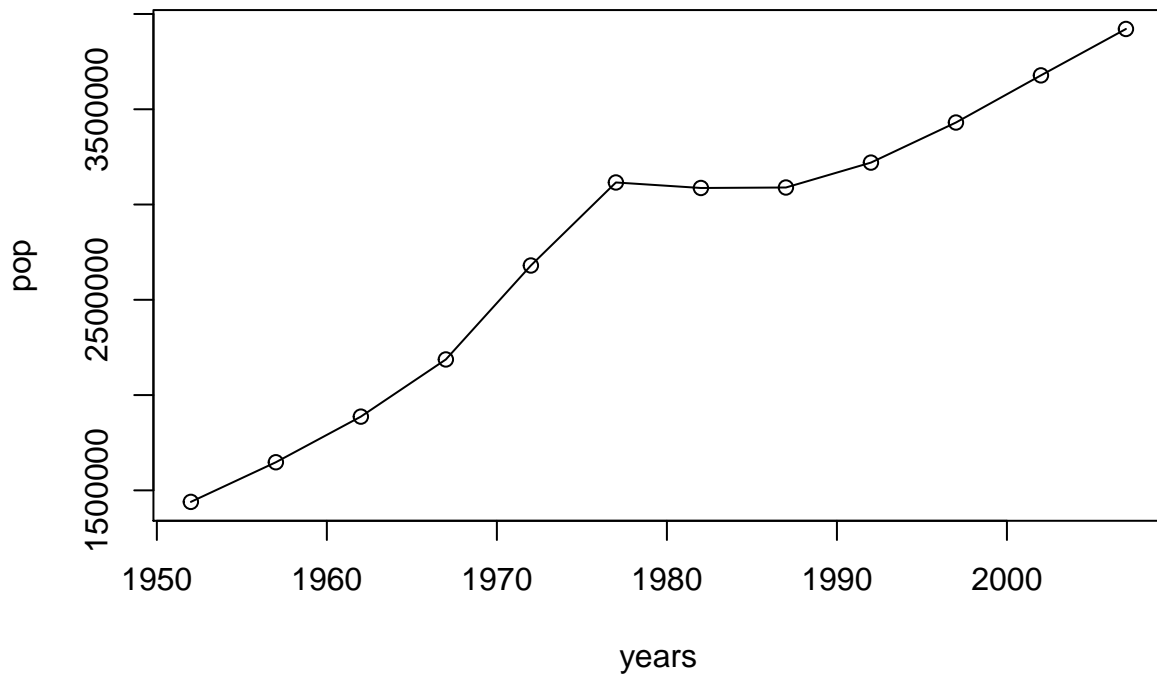


There

was a decrease from 1987 to 1992 in Kuwait.

For Lebanon:

Total population in Afghanistan from 1952 to 2007

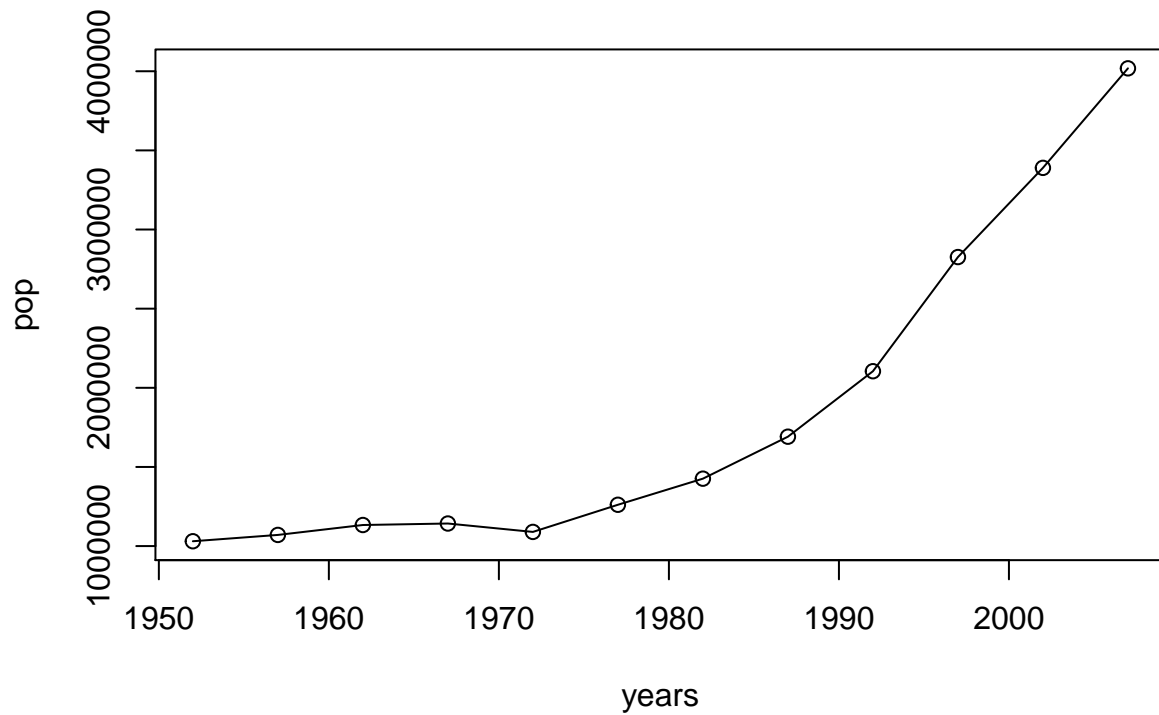


There

was a decrease from 1977 to 1987 in Lebanon.

For West Bank and Gaza:

Total population in Afghanistan from 1952 to 2007



There was a decrease from 1967 to 1972 in West Bank and Gaza.

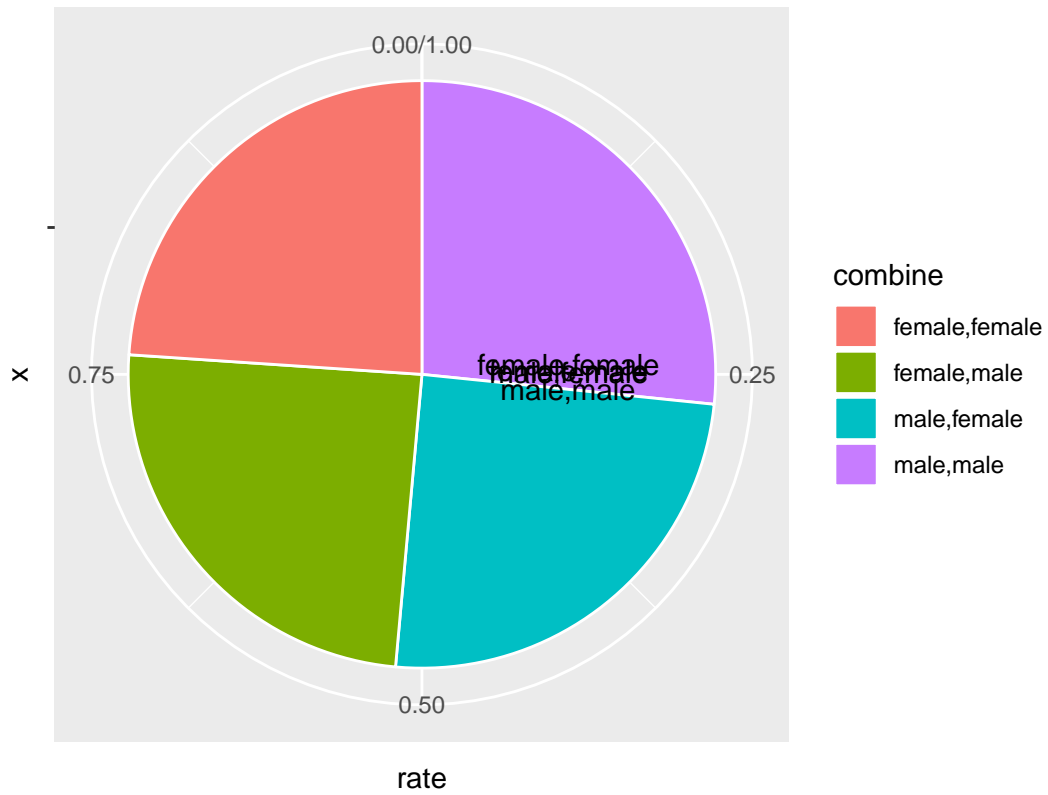
Which countries in the dataset have had the highest rate of growth in per capita GDP? Illustrate your answer with a table or plot.

Equatorial Guinea has the highest rate of growth in per capita GDP 3135.542%.

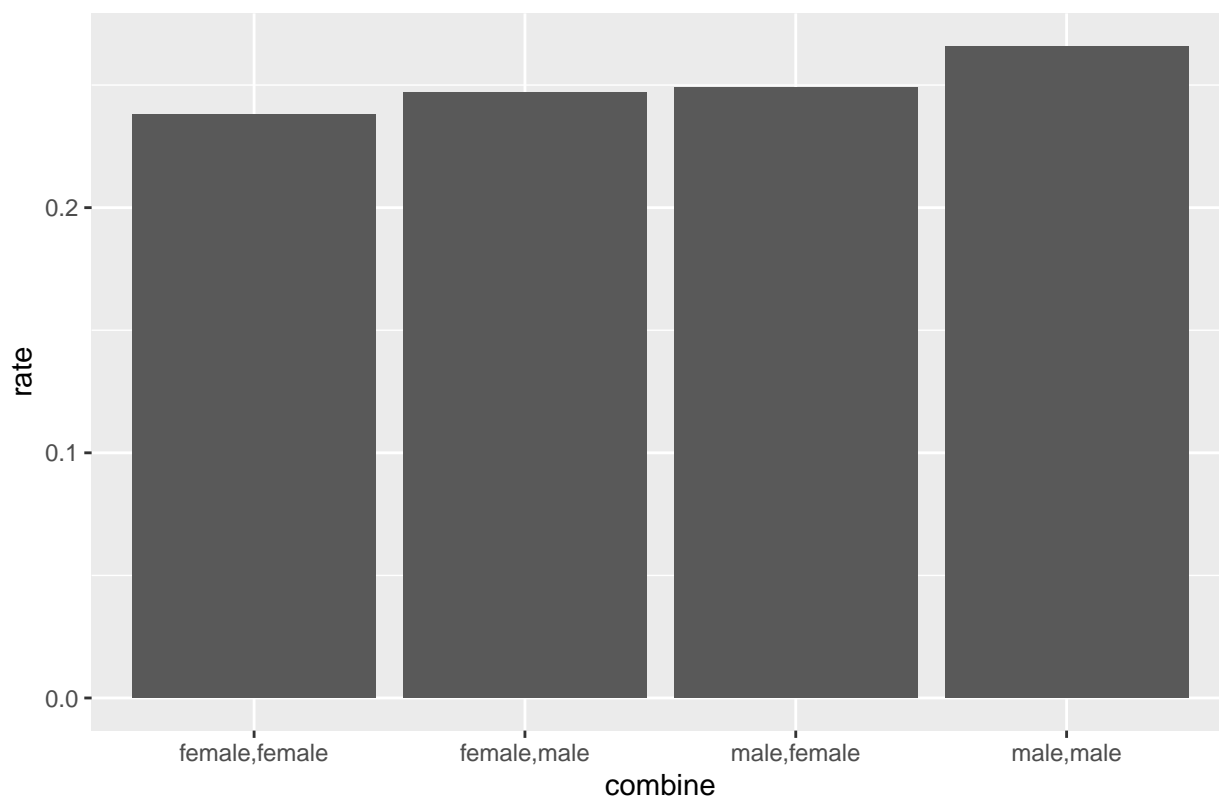
Problem 2

There are four possible gender combinations for the first two Children. Product a plot the contracts the frequency of these four combinations. Are the frequencies different for women in their 20s and women who are older than 29?

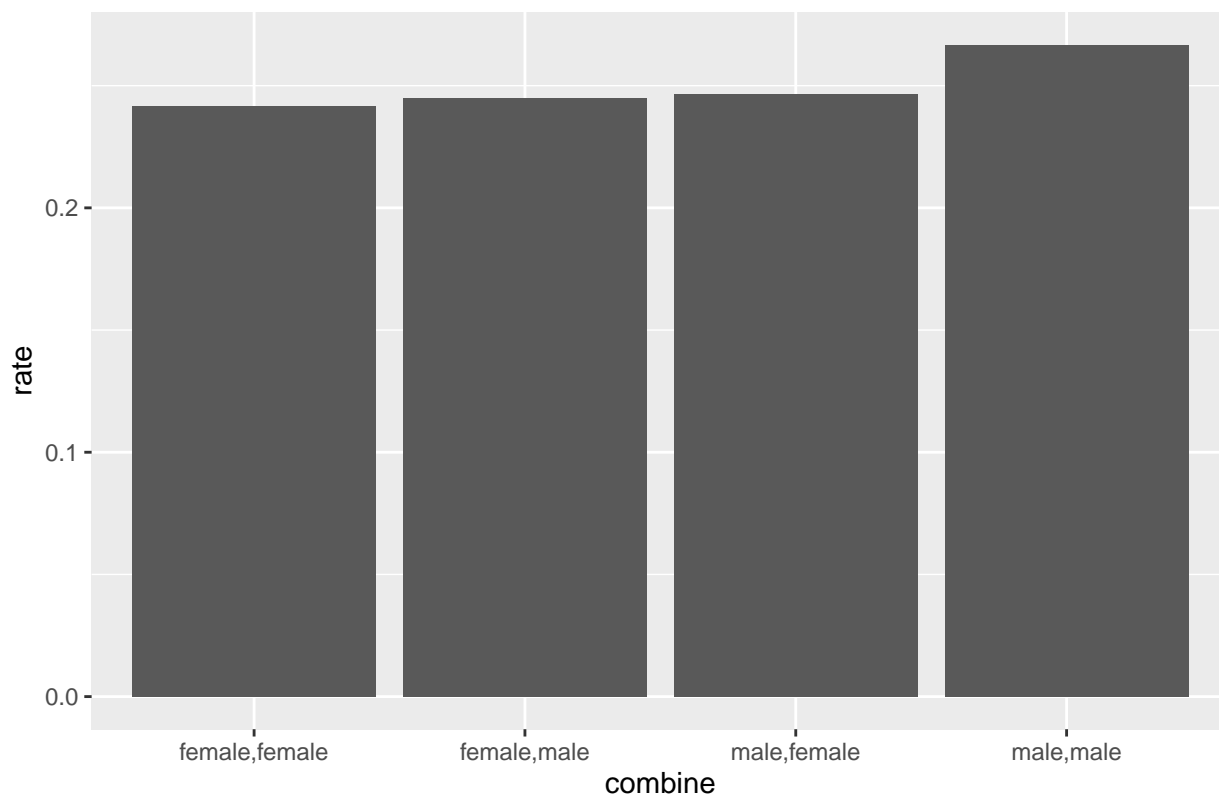
Frequency of four combinations



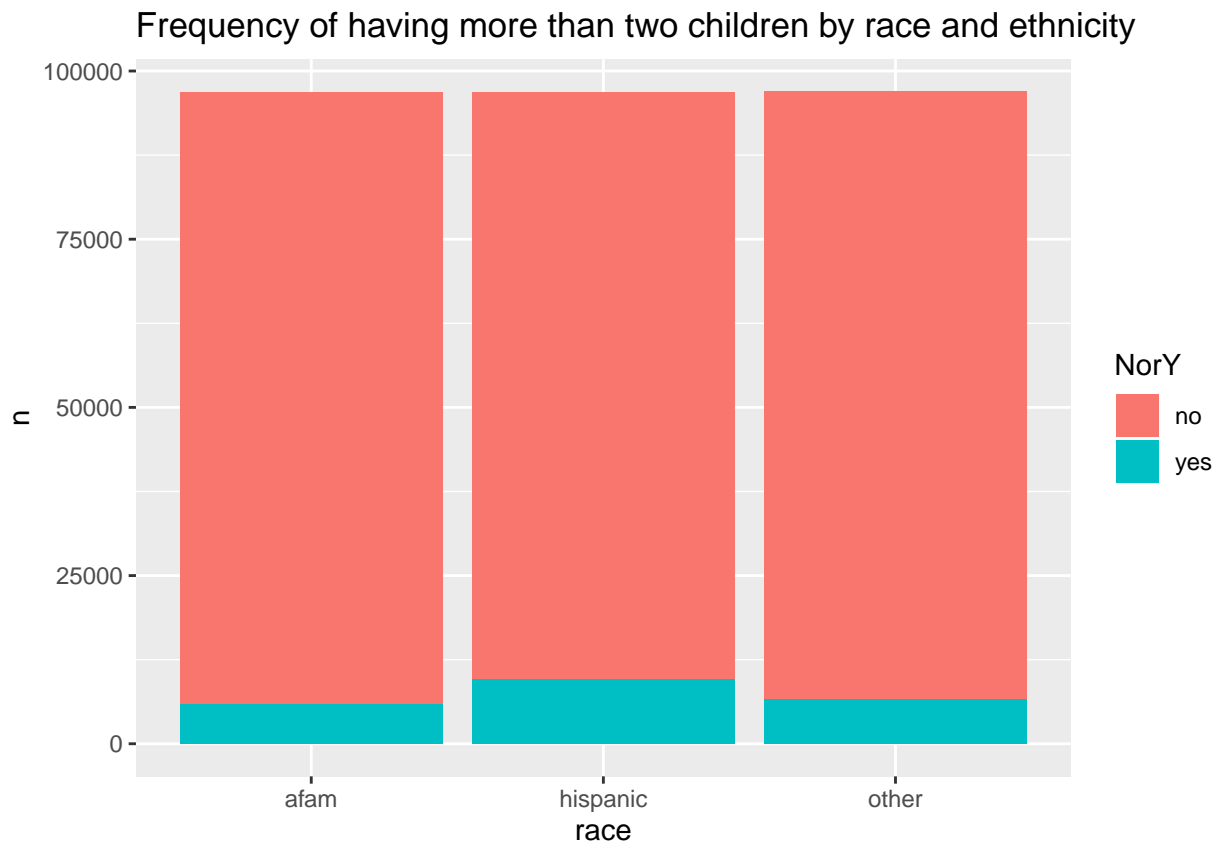
Frequency of four combinations for old women



Frequency of four combinations for young women



Produce a plot that contrasts the frequency of having more than two children by race and ethnicity.



Problem 3

Use the mtcars and mpg datasets.

How many times does the letter “e” occur in mtcars rownames?

There are 25 “e” occur in the mtcars rownames.

How many cars in mtcars have the brand Merc?

There are 7 cars in mtcars have the brand Merc.

How many cars in mpg have the brand (“manufacturer” in mpg) Merc?

There are 4 cars in mpg have the brand Merc.

Contrast the mileage data for Merc cars as reported in mtcars and mpg. Use tables, plots, and a short explanation.

Problem 4

Table 4: Mileage data for Merc cars in mtcars

manufacturer	mpg
Merc 240D	24.4
Merc 230	22.8
Merc 280	19.2
Merc 280C	17.8
Merc 450SE	16.4
Merc 450SL	17.3
Merc 450SLC	15.2

Table 5: Mileage data for Merc cars in mpg

manufacturer	cty	hwy
mercury	14	17
mercury	13	19
mercury	13	19
mercury	13	17

Produce a tabble that displays the five most popular boy names and girl names in the years 1880,1920, 1960, 2000.

What names overlap boys and girls?

In the sample size, there are 51500 names overlap boys and girls.

What names were used in the 19th century but have not been used in the 21th century?

There 1040 names were used in the 19th century but have not been used in the 21th century.

Table 6: The five most popular boy names and girl names in the years 1880,1920, 1960, 2000

1880		1920		1960		2000	
Male	Female	Male	Female	Male	Female	Male	Female
Charlie	Gertrude	Jack	Lois	Daniel	Sharon	Robert	Samantha
David	Jennie	Arthur	Lucille	Timothy	Nancy	Austin	Alexis
Frank	Carrie	Henry	Catherine	William	Donna	Alexander	Sarah
James	Cora	Paul	Louise	Robert	Karen	David	Hannah
John	Anna	James	Anna	Michael	Linda	Andrew	Emily

Table 7: Frequency of the names – Donald, Hilary, Hillary, Joe, Barrack

name	n	frequency
Donald	19383	1.53
Hilary	1317	0.08
Hillary	3534	0.07
Joe	23616	0.64

Produce a chart that shows the relative frequency of the names “Donald”, “Hilary”, “Hillary”, “Joe”, “Barrack”, over the years 1880 through 2017.

Frequency of the names -- Donald, Hilary, Hillary, Joe, Barrack

